引文格式：秦锋, 王涛, 张振虎. 空间直线拟合新方法[J]. 地理空间信息, 2023, 21(3): 21-24. doi:10. 3969/j.issn. 1672-4623.2023.03.005

空间直线拟合新方法

秦　锋[1]，王　涛[2]，张振虎[3]
（1. 中国核工业集团三门核电有限公司，浙江 台州 317112；2. 武汉市勘察设计有限公司，湖北 武汉 430030；
3. 上海勘察设计研究院（集团）有限公司，上海 200093）

摘　要：首先证明了正交距离最小二乘准则下拟合的空间直线必经过数据中心点，再以此为基础提出了一种空间直线拟合新算法。该方法不用迭代，计算简便。证明了正交距离最小二乘准则和整体最小二乘准则在拟合空间直线上的等价性；为增强算法的抗粗差能力，增加了RANSAC策略识别粗差；结合案例，分别采用BFGS、SQP优化算法以及整体最小二乘法对该方法进行验证，从而证明其准确性和有效性。

关键词：空间直线；拟合；正交距离；中心点；奇异值分解
中图分类号：P258　　**文献标志码：**B　　**文章编号：**1672-4623（2023）03-0021-04

Spatial Straight Line Fitting New Method

QIN Feng[1], WANG Tao[2], ZHANG Zhenhu[3]
(1. CNNC Sanmen Nuclear Power Co., Ltd., Taizhou 317112, China;
2. Wuhan Geotechnical Engineering and Surveying Co., Ltd., Wuhan 430030, China;
3. Shanghai Research Institute of Survey and Design(Group) Co., Ltd., Shanghai 200093, China)

Abstract: In this paper, we proved that the spatial straight line fitted under the orthogonal distance least squares criterion must passed through the center point of data at first. On the basis of this, we proposed a new spatial straight line fitting method. This method did not require iteration and was easy to calculate. We also proved the equivalence of orthogonal distance least squares criterion and the total least squares criterion in the spatial straight line fitting. At the same time, in order to enhance the algorithm ability to resist gross errors, we added the RANSAC strategy in the method to identify gross errors. Finally, combining with the case, we used BFGS, SQP optimization algorithm and the total least squares method to verify the method proposed in this paper, which could prove the accuracy and effectiveness of this method.

Key words: spatial straight line, fitting, orthogonal distance, center point, singular value decomposition

工程应用中经常遇到直线拟合问题。平面直线拟合通常采用最小二乘（LS）、整体最小二乘[1-2]（TLS）、正交距离拟合[3]以及奇异值分解法[4]。工程应用中不仅涉及平面直线拟合问题，还会遇到空间直线拟合问题。空间直线拟合比平面直线拟合复杂，目前的主要方法为TLS法和正交距离拟合法。采用TLS法拟合空间直线的研究较多，如杜明芳[5]基于TLS准则采用牛顿—梯度算法进行空间直线拟合；姚宜斌[6]等先对空间直线方程进行适当变换，再采用TLS法拟合空间直线；邓小渊[7]、汪奇生[8]等采用基于PEIV模型的结构总体最小二乘法（STLS）拟合空间直线。采用正交距离法拟合空间直线的研究较少，如曹慧荣[9]等基于正交距离最小二乘法和最小一乘法，分别采用主成分分析法和单纯形算法求解拟合空间直线参数；潘国荣[10]等则采用基于正交距离的以加权全最小二乘（WFLS）为准则的特征分解与选权迭代法拟合空间直线。上述方法大多需进行迭代，计算较复杂。

本文基于正交距离最小二乘准则提出了一种空间直线拟合新算法。该方法无需迭代，计算简便。同时，为探讨正交距离最小二乘法和TLS法在空间直线拟合上的联系，本文利用拉格朗日乘子法证明了两种方法在空间直线拟合上是等价的。为准确识别并剔除粗差，本文在新算法的基础上添加了基于随机采样一致性算法（RANSAC）[11]的粗差处理策略。结合案例，本文分别采用数值优化算法和STLS法验证了本文算法的准确性和有效性。

收稿日期：2021-12-10；修回日期：2022-03-18。
项目来源：上海市科学技术委员会科研计划资助项目（19DZ1201700）。
第一作者简介：秦锋（1985—），高级工程师，主要研究方向为核电工程测量及激光跟踪测量数据处理，E-mail：414453147@qq.com。

1 新方法模型与理论推导

1.1 空间直线的表现形式

空间直线一般可表示为：

$$\frac{x-X_0}{A}=\frac{y-Y_0}{B}=\frac{z-Z_0}{C} \tag{1}$$

由式（1）可知，确定一条空间直线需要空间直线的向量 $[A\ B\ C]$ 和直线上一点 (X_0, Y_0, Z_0)。对式（1）进行适当变换，并令 $a=\frac{A}{C}$，$b=\frac{B}{C}$，$x_0=X_0-aZ_0$，$y_0=Y_0-bZ_0$，则空间直线方程可表示为：

$$\frac{x-x_0}{a}=\frac{y-y_0}{b}=z \tag{2}$$

设拟合空间直线的三维点为 $(x_i, y_i, z_i), 1\leqslant i\leqslant n$，$(x_i, y_i, z_i)$ 到待拟合空间直线的几何距离为 d_i，即点到直线的正交距离。结合式（2），d_i 的计算公式为：

$$d_i=\frac{\sqrt{[b(x_i-x_0)-a(y_i-y_0)]^2+[(x_i-x_0)-az_i]^2+[(y_i-y_0)-bz_i]^2}}{\sqrt{a^2+b^2+1}} \tag{3}$$

1.2 正交距离最小二乘准则下拟合的空间直线经过数据中心点证明

以正交距离最小二乘为准则，构造空间直线拟合优化函数：

$$f_{\text{od}}=\sum_{i=1}^{n}d_i^2 \tag{4}$$

在 f_{od} 取得最小值时得到的空间直线参数为最佳参数，即求解无约束最小二乘优化问题 $\min f_{\text{od}}$。

根据极值条件可知，f_{od} 取极小值时其对参数 x_0、y_0 的偏导数应等于0，则有：

$$\begin{cases}\dfrac{\partial f_{\text{od}}}{\partial x_0}=\displaystyle\sum_{i=1}^{n}\frac{-2(b^2+1)(x_i-x_0)+2ab(y_i-y_0)+2az_i}{a^2+b^2+1}=0\\ \dfrac{\partial f_{\text{od}}}{\partial y_0}=\displaystyle\sum_{i=1}^{n}\frac{-2(a^2+1)(y_i-y_0)+2ab(x_i-x_0)+2bz_i}{a^2+b^2+1}=0\end{cases} \tag{5}$$

$$x_0=\sum_{i=1}^{n}\frac{x_i}{n}-a\sum_{i=1}^{n}\frac{z_i}{n}\quad y_0=\sum_{i=1}^{n}\frac{y_i}{n}-b\sum_{i=1}^{n}\frac{z_i}{n} \tag{6}$$

由式（6）可知，正交距离最小二乘准则下拟合的空间直线经过点（$\sum_{i=1}^{n}\frac{x_i}{n}-a\sum_{i=1}^{n}\frac{z_i}{n}, \sum_{i=1}^{n}\frac{y_i}{n}-b\sum_{i=1}^{n}\frac{z_i}{n}, 0$），将该点在经过该点且向量为 $[a\ b\ 1]$ 的空间直线上移动，使 Z 坐标为 $\sum_{i=1}^{n}\frac{z_i}{n}$，此时得到的 X、Y 坐标分别为 $\sum_{i=1}^{n}\frac{x_i}{n}$、$\sum_{i=1}^{n}\frac{y_i}{n}$，表明正交距离最小二乘准则下拟合的空间直线必经过数据中心点 $(\sum_{i=1}^{n}\frac{x_i}{n}, \sum_{i=1}^{n}\frac{y_i}{n}, \sum_{i=1}^{n}\frac{z_i}{n})$。

1.3 空间直线拟合新方法

§1.2节已证明了正交距离最小二乘准则下拟合的空间直线必过数据中心点。中心点坐标 (x_g, y_g, z_g) 可表示为：

$$x_g=\sum_{i=1}^{n}\frac{x_i}{n}, y_g=\sum_{i=1}^{n}\frac{y_i}{n}, z_g=\sum_{i=1}^{n}\frac{z_i}{n} \tag{7}$$

式（1）的空间直线方程可变为：

$$\frac{x-x_g}{A}=\frac{y-y_g}{B}=\frac{z-z_g}{C} \tag{8}$$

将式（8）同乘以 $\sqrt{A^2+B^2+C^2}$，并对数据进行中心化处理，令 $\frac{A}{\sqrt{A^2+B^2+C^2}}=j, \frac{B}{\sqrt{A^2+B^2+C^2}}=k, \frac{C}{\sqrt{A^2+B^2+C^2}}=l, \bar{x}=x-x_g, \bar{y}=y-y_g, \bar{z}=z-z_g$，则空间直线方程变为：

$$\frac{\bar{x}}{j}=\frac{\bar{y}}{k}=\frac{\bar{z}}{l}\quad \text{s.t. } j^2+k^2+l^2=1 \tag{9}$$

式中，$(\bar{x}, \bar{y}, \bar{z})$ 为中心化后的坐标；$[j\ k\ l]$ 为空间直线单位向量；s.t. 为 subject to 的缩写，表示受限制于某条件。

中心化处理后点到空间直线的距离可表示为：

$$\begin{gathered}\bar{d}_i=\sqrt{(k\bar{x}_i-j\bar{y}_i)^2+(l\bar{x}_i-j\bar{z}_i)^2+(l\bar{y}_i-k\bar{z}_i)^2}\\ \text{s.t. } j^2+k^2+l^2=1\end{gathered} \tag{10}$$

点到空间直线的距离平方和可表示为：

$$f_g=\sum_{i=1}^{n}\bar{d}_i^2\quad \text{s.t.} j^2+k^2+l^2=1 \tag{11}$$

f_g 在满足限制条件取最小值时得到的参数 j、k、l 为最佳参数，即求解以下约束优化问题：

$$\min f_g\quad \text{s.t.} j^2+k^2+l^2=1 \tag{12}$$

f_g 分别对 j、k、l 求偏导数并置0，则有：

$$\begin{cases}\dfrac{\partial f_g}{\partial j}=2j\displaystyle\sum_{i=1}^{n}\bar{y}_i^2-2k\sum_{i=1}^{n}\bar{x}_i\bar{y}_i+2j\sum_{i=1}^{n}\bar{z}_i^2-2l\sum_{i=1}^{n}\bar{x}_i\bar{z}_i=0\\ \dfrac{\partial f_g}{\partial k}=2k\displaystyle\sum_{i=1}^{n}\bar{x}_i^2-2j\sum_{i=1}^{n}\bar{x}_i\bar{y}_i+2k\sum_{i=1}^{n}\bar{z}_i^2-2l\sum_{i=1}^{n}\bar{y}_i\bar{z}_i=0\\ \dfrac{\partial f_g}{\partial l}=2l\displaystyle\sum_{i=1}^{n}\bar{x}_i^2-2j\sum_{i=1}^{n}\bar{x}_i\bar{z}_i+2l\sum_{i=1}^{n}\bar{y}_i^2-2k\sum_{i=1}^{n}\bar{y}_i\bar{z}_i=0\end{cases} \tag{13}$$

将上述3个等式以矩阵形式表示：

$$\begin{bmatrix}\displaystyle\sum_{i=1}^{n}(\bar{y}_i^2+\bar{z}_i^2) & -\displaystyle\sum_{i=1}^{n}\bar{x}_i\bar{y}_i & -\displaystyle\sum_{i=1}^{n}\bar{x}_i\bar{z}_i\\ -\displaystyle\sum_{i=1}^{n}\bar{x}_i\bar{y}_i & \displaystyle\sum_{i=1}^{n}(\bar{x}_i^2+\bar{z}_i^2) & -\displaystyle\sum_{i=1}^{n}\bar{y}_i\bar{z}_i\\ -\displaystyle\sum_{i=1}^{n}\bar{x}_i\bar{z}_i & -\displaystyle\sum_{i=1}^{n}\bar{y}_i\bar{z}_i & \displaystyle\sum_{i=1}^{n}(\bar{x}_i^2+\bar{y}_i^2)\end{bmatrix}\begin{bmatrix}j\\k\\l\end{bmatrix}=0 \tag{14}$$

令式（14）的系数矩阵为 $\boldsymbol{M}$，$\boldsymbol{X}=[j\ k\ l]^{\mathrm{T}}$，式（14）可转换为 $\boldsymbol{MX}=0$；且向量 $[j\ k\ l]$ 满足 $j^2+k^2+l^2=1$，因此有限制条件 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}=1$，即求解以下问题：

$$\boldsymbol{MX}=0 \quad \text{s.t.}\ \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}=1 \tag{15}$$

对 $\boldsymbol{M}$ 进行奇异值分解，即

$$\boldsymbol{M}=\boldsymbol{U}\Sigma\boldsymbol{V}^{\mathrm{T}} \tag{16}$$

式中，$\boldsymbol{U}=[U_1\ \ U_2\ \ U_3]$ 为左奇异矩阵；$\boldsymbol{V}=[V_1\ \ V_2\ \ V_3]$ 为右奇异矩阵；$\Sigma=\mathrm{diag}(\sigma_1,\sigma_2,\sigma_3)$，$\sigma_1$、$\sigma_2$、$\sigma_3$ 为 $\boldsymbol{M}$ 的奇异值。

根据参考文献[4]，最小奇异值 σ_3 对应的右奇异向量 $\boldsymbol{V}_3$ 即为所求的空间直线单位向量 $[j\ k\ l]^{\mathrm{T}}$。

1.4　粗差识别与处理

实际工程应用中，测量数据采集过程难免出现粗差，为提高参数计算准确度，有必要在采用上述算法计算空间直线拟合参数前识别并剔除粗差。本文采用RANSAC算法识别和剔除粗差点。RANSAC算法的主要思想是随机采用较少的点估算出初始模型，然后采用剩下的点验证模型，从中选取一组使最多数量的点表现良好的模型。具体到空间直线拟合，利用RANSAC算法的步骤为：①从原始数据中随机选取两个点连接成空间直线，计算剩余点到空间直线的距离；②设置阈值 η，若点到空间直线的距离小于阈值，则视为正常点，否则视为异常点；③重复步骤①和②，选取正常点最多的一组模型。选取正常点最多的一组模型后，可认为剩余点为异常点，即含有粗差的点。剔除粗差点后，即可采用§1.3节的算法进行空间直线拟合。

2　正交距离最小二乘法与TLS法拟合空间直线等价性证明

假设 (x_i,y_i,z_i) 中存在误差 (ex_i,ey_i,ez_i)，且误差独立同分布，根据式（2）建立以下变量含误差模型，即

$$\begin{cases} x_i+ex_i=a(z_i+ez_i)+x_0 \\ y_i+ey_i=b(z_i+ez_i)+y_0 \end{cases} \tag{17}$$

式中，$1\leqslant i\leqslant n$，下同。

基于TLS准则，构造优化函数，即

$$f_{\mathrm{TLS}}=\sum_{i=1}^{n}(ex_i^2+ey_i^2+ez_i^2) \tag{18}$$

求解最佳拟合参数就转变为以下约束最优化问题：

$$\begin{aligned} &\min f_{\mathrm{TLS}} \quad \text{s.t.}\ x_i+ex_i=a(z_i+ez_i)+x_0 \\ &\text{s.t.}\ y_i+ey_i=b(z_i+ez_i)+y_0 \end{aligned} \tag{19}$$

该约束优化问题可采用拉格朗日乘子法予以解决。首先构造拉格朗日函数：

$$\begin{aligned} f_{LA}=&\sum_{i=1}^{n}(ex_i^2+ey_i^2+ez_i^2)-2\sum_{i=1}^{n}\mu_{2i-1}(x_i+\Delta x_i-az_i-a\Delta z_i- \\ &x_0)-2\sum_{i=1}^{n}\mu_{2i}(y_i+\Delta y_i-bz_i-b\Delta z_i-y_0) \end{aligned} \tag{20}$$

对 ex_i、ey_i、ez_i、μ_{2i-1} 和 μ_{2i} 分别求偏导数并置0，可得：

$$\begin{cases} 2\Delta x_i=2\mu_{2i-1} \\ 2\Delta y_i=2\mu_{2i} \\ 2\Delta z_i+2a\mu_{2i-1}+2b\mu_{2i}=0 \end{cases} \tag{21}$$

由上述等式可解算得到：

$$\begin{cases} \Delta x_i=\dfrac{(b^2+1)(x_0-x_i)-ab(y_0-y_i)+az_i}{a^2+b^2+1} \\ \Delta y_i=\dfrac{(a^2+1)(y_0-y_i)-ab(x_0-x_i)+bz_i}{a^2+b^2+1} \\ \Delta z_i=\dfrac{-a(x_0-x_i)-b(y_0-y_i)-(a^2+b^2)z_i}{a^2+b^2+1} \end{cases} \tag{22}$$

将式（22）代入 f_{LA} 中，可得：

$$f_{\mathrm{LA}}=f_{\mathrm{od}} \tag{23}$$

从而证明正交距离最小二乘法与TLS法在拟合空间直线上是等价的。

3　案例分析

本文采用参考文献[12]数据，原始数据如表1所示，已知表中模型测量误差 $\delta=\pm0.005$。

表1　原始数据

序号	X	Y	Z
1	3.003 6	2.996	3.004 1
2	4.003 4	4.998	6.003 3
3	5.005	6.999 2	9.005
4	5.996 4	9.003 6	11.996 2
5	7.003 2	10.996 8	15.002 4
6	8.003 8	13.000 1	18.004 5
7	9.001	15.004	20.999 6
8	9.997 2	16.999 3	23.998 2
9	11.001 1	18.996 2	27.000 1
10	11.999 7	20.998	29.997 8

3.1　参数计算

本文分别采用本文算法和STLS法对上述数据进行处理，并以两种数值优化算法结果作为验证。对于式（4）中的正交距离最小二乘拟合空间直线问题，采用Broyden–Fletcher–Goldfarb–Shanno（BFGS）算法进行验证求解；对于式（19）中的TLS拟合空间直线问题，采用序列二次规划法（SQP）进行验证求解。BFGS算法和SQP算法详见参考文献[13]。BFGS算法、SQP算法以及STLS法均需进行迭代计算，根据第一个点和第二个点计算空间直线拟合参数初始值。迭代终止条件为两次拟合参数之差的二范数 $\leqslant 10^{-10}$。各算法

的计算结果如表2所示，可以看出，本文算法无需计算初始值，也无需迭代；而BFGS算法、SQP算法和STLS法分别需要迭代8次、3次和10次，因此本文算法计算效率优于其他算法。

表2 各算法计算结果

参数	本文算法	BFGS算法	SQP算法	STLS法
a	0.267 179 825 5	0.267 179 825 5	0.267 179 825 5	0.333 248 094 9（0.267 179 825 5）
b	0.534 622 021 3	0.534 622 021 3	0.534 622 021 3	0.666 823 439 0（0.534 622 021 3）
c	0.801 744 494 9	0.801 744 494 9	0.801 744 494 9	1（0.801 744 494 9）
x_0	7.501 44（2.002 473 196 6）	2.002 473 196 6	2.002 473 196 6	2.002 473 196 6
y_0	11.999 12（0.995 786 413 4）	0.995 786 413 4	0.995 786 413 4	0.995 786 413 5
z_0	16.501 12（0）	0	0	0
迭代次数	无迭代	8	3	10

注：为便于数据对比，将STLS法计算得到的空间直线向量单位化，括号内为单位化后的向量；将本文算法得到的空间直线上一点的坐标移动至$z_0=0$，括号内为移动后的数据。

3.2 计算结果分析与精度统计

对比各算法的计算结果发现，本文算法与BFGS算法、SQP算法和STLS法的计算结果完全一致，说明本文算法准确可靠，同时验证了本文算法与TLS法的等价性。对于空间直线的拟合精度，可通过计算各拟合点到拟合直线的距离平方和、距离之和以及直线度等评定。各点到拟合直线的距离可通过式（3）计算。直线度的计算公式为：

$$p=\max(d_i)-\min(d_m) \tag{24}$$

本文计算了各算法拟合的空间直线的直线度、距离之和以及距离平方和，由于各算法计算结果一致，因此只展示本文算法的计算结果：$\max(d_i)=0.008\ 581$，$\min(d_m)=0.000\ 900$，$p=0.007\ 681$，$\sum_{i=1}^{n}d_i=0.030\ 727$，$\sum_{i=1}^{n}d_i^2=0.000\ 138$。

3.3 增加粗差后的参数计算

对8~10号点添加10倍中误差作为粗差，再采用本文算法计算空间直线拟合参数，即先采用RANSAC方法识别并剔除粗差，再采用§1.3节中的方法计算拟合参数。经试验，上述方法可有效剔除全部粗差点，剔除粗差点后拟合的空间直线单位向量为[0.267 176 0.534 766 0.801 649]，直线上一点的坐标为（2.002 228, 0.993 243,0）。对比表2中的计算结果发现，两次计算的拟合参数非常接近，说明本文算法是有效的。

4 结 语

1）本文首先证明了正交距离最小二乘准则下拟合的空间直线必经过数据中心点，并以此为基础，提出了一种空间直线拟合新方法。该方法简单实用，无需计算初始值和迭代运算，采用数据中心化和奇异值分解法即可求解空间直线拟合参数。

2）本文证明了正交距离最小二乘法和TLS法在拟合空间直线上的等价性，并结合案例，采用数值优化算法和其他算法进行了验证。

3）为增强算法的抗粗差能力，本文采用了RANSAC粗差处理策略。经实验验证，该策略可有效识别并剔除粗差。

参考文献

[1] 鲁铁定,陶本藻,周世健. 基于整体最小二乘法的线性回归建模和解法[J]. 武汉大学学报(信息科学版),2008,33(5):504-507

[2] 丁克良,沈云中,欧吉坤. 整体最小二乘法直线拟合[J]. 辽宁工程技术大学学报(自然科学版),2010,29(1):44-47

[3] 崔佳洁,崔希民,王强,等. 基于最小二乘的正交距离拟合方法研究[J]. 计算机工程与应用,2015,51(11):27-30

[4] 张贤达. 矩阵分析与应用(第2版)[M]. 北京:清华大学出版社,2013

[5] 杜明芳. 空间直线拟合[J]. 北京印刷学院学报,1996,4(2):27-31

[6] 姚宜斌,黄书华,孔建,等. 空间直线拟合的整体最小二乘算法[J]. 武汉大学学报(信息科学版),2014,39(5):571-574

[7] 邓小渊,鲁铁定,常晓涛,等. 空间直线拟合的混合结构总体最小二乘算法[J]. 测绘科学,2016,41(6):1-4

[8] 汪奇生,杨德宏,杨腾飞. 空间直线的结构总体最小二乘拟合[J]. 大地测量与地球动力学,2015,35(3):433-435

[9] 曹慧荣,王福昌. 全最小二乘和全加权最小一乘准则下的空间直线拟合[J]. 计算机应用与软件,2010,27(9):111-112,127

[10] 潘国荣,唐杭. 特征分解与选权迭代在空间直线拟合中的应用[J]. 东南大学学报(自然科学版),2013,43(增刊2):250-255

[11] Fischler M A, Bolles R C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography[J]. Readings in Computer Vision,1981,24(6):381-395

[12] 陈基伟. 工业测量数据拟合研究[D]. 上海:同济大学,2005

[13] 袁亚湘,孙文瑜. 最优化理论与方法[M]. 北京:科学出版社,2007